# Extending and Applying the EPIC Architecture for Human Cognition and Performance

## Final Report
## Project N00014-06-1-0034

David E. Kieras
University of Michigan

## Report No. FR-11/ONR-EPIC-17

## Period Covered: 1 October 2005 – 30 September 2010


Requests for copies should be sent to: David E. Kieras, Electrical Engineering & Computer Science Department, University of Michigan, 3641 Computer Science & Engineering Building, 2260 Hayward Street, Ann Arbor, MI 48109-2121, kieras@umich.edu.

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 074-0188*



| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | 20 January 2011 | Final Report 1 October 2005 – 30 September 2010 |

**4. TITLE AND SUBTITLE**
Extending and Applying the EPIC Architecture for Human Cognition and Performance

**5. FUNDING NUMBERS**
N00014-06-1-0034
09PR01964-02

**6. AUTHOR(S)**
David E. Kieras

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
University of Michigan Division of Research Development and Administration, Ann Arbor, MI 48109

**8. PERFORMING ORGANIZATION REPORT NUMBER**
FR-11/ONR-17

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
Office of Naval Research (Code 342)
875 N. Randolph St.
Arlington, VA 22217-5660

**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for Public Release; Distribution Unlimited

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 Words)*
This is the final report for a project that was in a series of projects on the development and validation of the EPIC cognitive architecture for modeling human cognition and performance. The report summarizes the results and lists the products and publications resulting from the project. These including the modeling system software, fundamental architecture development, modeling of visual search, and complex task performance.

**14. SUBJECT TERMS**
Cognitive Architecture, Human Performance Modeling

**15. NUMBER OF PAGES**
18

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| UNCLASSIFIED | UNCLASSIFIED | UNCLASSIFIED | UL |

# Extending and Applying the EPIC Architecture for Human Cognition and Performance

## ONR Grant N00014-06-1-0034

## Period Covered: 1 October 2005 – 30 September 2010

David Kieras, Principal Investigator

## Summary of Project Goals and Accomplishments

This is the final report for a project on the development and validation of the EPIC cognitive architecture for modeling human cognition and performance. It continued a series of ONR-sponsored projects on the development of the EPIC architecture for human cognition and performance, conducted by the same PI and also David Meyer of the University of Michigan. This extended activity produced a large number of products and accomplishments; however, this report documents the outcome of only this specific project.

In this project, the EPIC computational architecture for modeling human cognition and performance was extended and applied to tasks especially relevant to military applications. The project took advantage of other Navy-sponsored research on modeling complex tasks by comparing EPIC models to models built with the lower-fidelity but more usable GLEAN architecture for GOMS modeling, and combining the best features of both architectures. This project contributed to the developing capability of human performance modeling to help design future human-machine systems to be maximally effective.

In the following section, each specific project goal is stated below, followed by a summary of the work accomplished for that goal.

### *Goal 1. The EPIC architecture components for motor control, audition, and speech will be revised and expanded.*

#### Visual search

The original proposal called for simply completing work on visual search already underway in the previous project, and then moving on to some other topics. However, it became clear that providing an adequate account of visual search was in fact an extremely challenging project all by itself. Very few cognitive modeling projects have attempted to account for such a span of related tasks at such a level of empirical detail. However, given the increasing reliance on complex visual displays to support military task performance, it is critical to have a fundamental understanding of how such displays are used.

The plan was to model visual search in several representative tasks that between them called into question the conventional wisdom about visual search and converged on important architectural issues. Findlay's (1997) data contradicted the conventional belief that visual search is guided by single perceptual features, and emphasized the speed of the process by which the target of a fixation could be chosen, which imposes stringent requirements on the architecture.

Peterson et al. (2001) contradicted the common belief that visual search has no memory by showing a highly reliable memory for fixations of at least 12 objects. The architecture has to accommodate such an effect without an ad-hoc memory mechanism just to record fixations. Williams (1967) showed that visual search of fields containing a large number of multidimensional objects is highly efficient and relies heavily on color to guide the search, at a scale beyond almost all other experiments. These results require that the architecture scale well, and the visual guidance has to be sensitive to the size of objects as well as their shape and color. Finally, the Marshall & St. John (St. John, et al. 2006) task involves realistically complex and numerous stimuli that differ in density, and show the same effects of multidimensional visual guidance and memory for fixations. The complete fixation sequence data were available, making it possible to stress the ability of the architecture to represent visual search processes at an unprecedented level of detail. Attempting to account for the data in all four of these tasks within a single architecture and consistent parameters and task strategies was a major challenge, much more than anticipated.

However, the basic result is that visual search can be modeled with the EPIC architecture making use of:

1. Probabilistic quadratic acuity functions that describe what visual features of an object are available in peripheral vision and can be used to guide the choice of the next object to fixate;

2. A persistent visual perceptual store that accumulates the information provided by successive fixations, building up a complete representation of the visual scene; an object can be chosen for fixation based on whether its representation in the store is complete;

3. A single basic task strategy for using the available information to guide the choice of the next fixation target, using overlapping processes for speed. Optimal use of the information appears to be a good first approximation.

The results of the modeling work were applied to improve the accuracy of the GOMS visual search operator in the GLEAN architecture.

This work has been presented in conference presentations and a journal article; further publications will be submitted.

### Motor control

An extensive literature review on the theoretical motor-control literature was conducted and integrated with recent key results from neuroscience. The major result is the need to incorporate an action system into the cognitive architecture to represent the set of brain structures associated with the posterior parietal cortex, also known as the "dorsal" system (e.g., Rossetti & Pisella, 2002). The action system is responsible for real-time coordination and control of visually-guided and spatially-situated movement, and is distinct from the cognitive system associated with conscious visual recognition and memory, the "ventral" system.

EPIC already has a suggestion of this basic architecture in that the motor subsystems for aimed manual movements and eye movements accept commands that specify the movement in terms of visual objects and then access the required object location and size information directly from the visual perceptual system. However, at present no mainstream computational cognitive architecture has this "dual CPU" sort of structure, meaning that fundamental phenomena involving spatially-situated visual objects and movements, such as typical radar workstation tasks, are incorrectly represented in the architecture.

These considerations led to a paper reviewing the motivation behind key features of the EPIC motor processors for aimed manual and eye movements, and motivating important modifications to them based on a reassessment of the motor control literature. Since other architectures such as ACT-R have adopted EPIC's characterization of the motor system, it was important to "go public" with the rationale for these changes; the paper is a contribution not just to computational cognitive architecture theory, but is also to the topic of motor control in human performance.

Current cognitive architectures lack an explicit action system, and thus distort fundamental aspects of human cognition and performance. Incorporating an action system will improve the scope and accuracy of computational cognitive architectures, especially for use in design problems where a wide range of human activities is involved.

### Audition and speech

Progress on these topics was made in the form of beginning a collaboration with Greg Wakefield, a colleague in the University of Michigan Electrical Engineering and Computer Science department, who is an expert on auditory signal processing and spatial audition phenomena. This led to the current project on developing the auditory and speech components of the EPIC architecture, with a focus on accounting for multichannel speech comprehension.

### *Goal 2. The EPIC and GLEAN software will modified to maximize code and model reuse between the two systems, and useful features of each added to the other.*

The planned software work to merge the code base for GLEAN and EPIC and enable device models to be interoperable between the two architectures received substantial attention and was completed. The code was updated to use the gcc C++ compiler and Mac OS X's GUI facilities, and various design improvements were made that will improve ease of maintenance and portability. The two architectures share many software components such as a standard visualization interface that shows what the simulated human can see, and now the very same code can now be used to create a simulated task environment for either EPIC or GLEAN models.

Many of the obstacles facing practical use of cognitive modeling are basically software problems, not just with the usability of the tools, but also the power and capability to scale to full-size problems and the compatibility of the modeling tools with the software development process (see Kieras, in press, for more discussion). The work with EPIC and GLEAN has been consistently focused on achieving modeling tools that can handle full-size problems, be easily maintained and extended as the architectures develop, and able to contribute to the system design process when not even a prototype has been developed. It is important to have these practical software problems being addressed along with the purely scientific concerns in cognitive architecture development.

The EPIC and GLEAN software have been distributed to additional users, and a Subversion source code repository website was set up to facilitate collaboration on EPIC development and distribution of the current version.

Work with MITRE on large-scale modeling of air traffic control tasks with GLEAN was an excellent proving ground for the GLEAN simulation software; MITRE used GLEAN for much more complex and large simulations than ever done by similar architectures, and it performed extremely well, both in raw computational power and relative ease of use. This is very encouraging since EPIC uses the same basic software approach.

The EPIC software package was modified to support modeling situations studied by Derek Brock's group at NRL and Anthony Hornof at University of Oregon, in which visual objects appear and disappear as a result of eye movements that bring them into or out of view, either because they are large eye/head movements as in Brock's dual-task models, or in a gaze-contingent paradigm as in Hornof's experiments on spatialized audio. Hornof's group also developed a temporal processing module for EPIC, based on the components in ACT-R. This is now available in the code repository.

Also to support this same work, EPIC's visual perceptual processor now supports dynamically loading a custom encoding function; as this capability is further developed, it will greatly simplify constructing models that require specialized perceptual recognition functions with the need to modify the architectural code.

A few tasks remain under the purely software work topic, but have been deferred while more substantive work was undertaken.

### *Goal 3. New capabilities for task working memory and decision making will be developed in EPIC models and added to the architecture.*

The plan was to address task working memory in the form of the Ericsson and Kintsch (1995) concept of long-term working memory. An opportunity appeared as a collaboration with Steven Estes of MITRE, who had been modeling working memory in air traffic control (ATC) and piloting tasks with informal GOMS models. He and his team began using the GLEAN tool to simulate human operators in their large simulations of ATC system, and we discussed both the relevant literature and some specific experimental plans to study working memory issues. One of these studies was to see if controllers had task-specific skilled memory analogous to chess players. The surprise was that controllers had very little post-session memory for any of the events in either typical or non-typical scenarios, unlike long-term working memory effects. Thus perhaps task working memory is not necessarily this form of working memory, but nonetheless it will be useful to further study the role of working memory in this complex real-world task.

The conclusion was that work on long-term working memory and decision procedures would have to be done in the context of Goal 4.

### *Goal 4. Apply EPIC models to a complex Navy-relevant task, and use GLEAN models to help inform development of the EPIC models.*

#### Revisiting the Muiti-Modal Watch Station Anti-Air Warfare task

A more accurate version of the Multi-Modal Watch Station Anti-Air Warfare task previously modeled with GLEAN was implemented in the EPIC software package which also provided a valuable built-in task visualization capability.

The data set collected under the SC-21 Multi-Modal Watch Station Anti-Air Warfare project (Osga, et al., 2002) was reassessed. There were some discrepancies in the device model and scenario used in the previous GLEAN models of teams constructed and validated by Santoro, Kieras, and Pharmer (2003). Even though the GLEAN models were based on the specified rules of engagement, their predictions were sometimes seriously off in ways that suggested that the human teams were not following the rules of engagement, meaning that it was not useful to attempt to model this data. However, perhaps some of the prediction error was actually due to these discrepancies in the device and scenario. Since data from complex and realistic Navy tasks, performed by real Navy personnel, is difficult and expensive to collect, it was worthwhile to more carefully determine whether the data could indeed support a serious model validation.

Using an EPIC version of the task environment, the accuracy of the device model and the scenario specifications were carefully checked; the EPIC visualization made this considerably easier than was originally possible with the GLEAN tool. Several discrepancies were uncovered and corrected, but the main value was clarifying how the task performers followed the rules of engagement in task execution: Some of the triggering events for tasks are simple, and these were well-predicted. However, more complex tasks do not have simple triggers, and the SME solution gives only time boundaries for these actions (no-earlier-than, no-later-than), which often either have no clear display events as triggers, or they depend on anticipating future display events. Comparing the model predictions to the data separately for these cases would allow a more accurate assessment of how the model can account for the data.

The original task scenario used in the data collection had an expert solution that was used as the basis of programming the model, but a detailed comparison of the empirical team data to the expert solution had not be previously made. This comparison was performed, and the results were that despite some systematic differences, the team performance was overall in accord with the expert solution, unlike the earlier impression. Thus the human team data is basically sound, despite its very limited sample size and lack of experimental controls, and can be used within its limits to validate models.

This task is a good context in which to examine some key cognitive architecture and task strategy issues. Although the data quality is poor by scientific standards due to the single scenario, small sample size, data collection limitations, and uncontrolled idiosyncratic team procedures, it still represents a rare and valuable resource on performance on a real Navy task that can be readily modeled. The limited quality of the data can be dealt with in terms of a criterion that might seem loose but which experience shows can be very hard to meet: If the model matches either the expert solution, or the typical human performance, in that it takes the same actions within the observed time brackets, then it is realistic and accurate enough to be useful in design decisions.

Beyond the demonstration of models for a complex task, this work contributes to the methodology of validating models of complex human activity, which is required for the sound application of human performance models to design problems. The methodological lessons from this (and other) work in the 2006 AFOSR workshop on Model Comparison and Validation.

### EPIC vs GLEAN models for the AAW task

Major progress was made in modeling the MMWS AAW task in EPIC, and using the model to begin answering some basic questions about architectural and strategy issues in such tasks. The work began by starting with the GLEAN GOMSL model for the task role responsible for monitoring the air track situation and issuing warnings and queries according to a set of Rules of Engagement (ROE), and writing the same GOMS methods as a set of hierarchically organized production rules. However, the effect of the differences in the GLEAN and EPIC architectures were immediately apparent, and have important implications for how these tasks should be modeled not just in EPIC, but in any architecture. Three important issues will be described.

*1. Maintaining the visual representation.* The GLEAN architecture "black boxes" visual search with a Mental operator that takes a moderate and fixed amount of time to identify an object meeting the search specifications. While this greatly simplifies model construction, it also ignores a variety of visual issues as well as some significant issues of cognitive strategy in such tasks. The EPIC model for the MMWS task started with the acuity functions developed in the modeling of the Marshall-St. John visual search task (see Kieras & Marshall, 2006) as a first approximation. This meant that the complete visual information was available for only a few screen objects at a time. The persistent visual store demonstrated in the visual search modeling helps ensure that more information is available, but some of the activities in the task, such as performing a query or examining the track data table, take the eyes from the main display for many seconds at a time and so would lead to a loss of visual information about the main display. Thus critical visual events could be missed before the missing information could be reacquired. It quickly emerged that the model was seriously dysfunctional as a result.

A better approach is that the task strategy includes some kind of "scanning" of the display to keep the perceptual store current, or refreshed, not unlike the standard airplane pilot practice of periodically scanning the cockpit instruments. A simple first approximation was built into the EPIC model: A background process picks a relevant object at random and fixates it, then repeats. This process is suspended whenever a relevant track event occurs, and is resumed whenever event handling is complete. Because EPIC supports native multithreaded processing, this process is simple to implement in production rules. The scanning strategy together with EPIC's large-capacity and persistent visual store means that cognition has available a fairly complete representation of the current visual situation at all times.

*2. Noticing task events.* The task involves performing actions in response to visual events. For example, according to the ROE, if a suspect track (e.g. a red triangle blip) crosses the Ownship-40nm range circle over international waters, the operator needs to issue a warning over a radio channel. The ROE defines several event types, and many events could occur close together in time. How is the search for relevant track events organized?

The GLEAN system shares with some other cognitive architectures a concept in which the visual system is a "black box" that is controlled top-down by central cognition – the visual system only locates objects as commanded by the task strategy. Thus, the task strategy "polls" the visual system for each possibility, serially in priority order. Continuing the above example, the GLEAN model would apply the GOMS visual search operator to determine if there is a red blip in international airspace inside the smaller of the two circles around the ownship icon; if not, it would apply the operator again with a different specification to determine if there is a red blip crossing the 75nm range circle around the CV in the task force; if not, it would apply the operator yet again to look for a yellow blip inside the 60nm range circle from ownship, and so forth. The simple GOMS model underlying GLEAN assigns a constant and relatively large amount of time to each such visual search operator. According to this model, many seconds would be required to perform all of these visual checks, even if only the last one is relevant.

While simple, the resulting model is quite counterintuitive and seriously inaccurate. Surely the human operator can look for these different kinds of events simultaneously, and can notice things other than what they might be deliberately searching for! This fact identifies a serious limitation of GLEAN – the tradeoff of accuracy in favor of simplicity has gone too far. The same problem would be suffered by any architecture in which cognition follows a strict goal-structured sequential procedure and attempts to "pull" information from the visual system, as in at least earlier forms of ACT-R.

Rather, a good model for such a task must allow parallel testing for possibilities in the visual scene and some kind of interruptability to allow events to be handled in order of their correct priority. Thus rather than search for visual events, the model must notice visual events. This requires that the architecture and model maintain a representation in which all (or at least most) relevant events are automatically represented (as described above). Then the model must monitor the representation and detect whatever events are occurring and select which to act on. The solution is to take advantage of EPIC's inherent capacity for true multithreading: a set of "daemon" rules wait in parallel for relevant events, following the nominate-choose process that has proved useful in the visual search models. The nomination rules are triggered by the different visual properties of screen objects and nominate events for consideration. The choose rules then pick the highest-priority nomination for processing. The background scanning process ensures that the nomination rules have current information to work from, so that within a short time, an event will be noticed, any background or lower priority activity gets suspended, and handling the new event is started.

The scanning and event noticing processes implement a simple form of visual situational awareness; by providing a clear and simple implementation in a plausible cognitive architecture, this work is a step forward in the understanding of this complex concept.

*3. What has already been done?* Many of actions done for a track are very time-consuming – requiring many seconds, and in some cases, such as issuing a query, should not be repeated haphazardly. How does the human operator keep track of what he or she has already done? The original AEGIS watch station setup apparently relied on devices outside the display for task support, such as paper notes or grease pencil on the CRT. The design philosophy of the MMWS was that it is not supposed to need such crutches; the multiple displays and task-oriented design should support the task fully. However, the models, both GLEAN and EPIC, of the basic first version of the MMWS design immediately show that the task apparently cannot be done in a purely display-based fashion; some form of memory seems to be required elsewhere. That is, the state of the display does not contain any information on what actions have been done – e.g. whether a track has been queried or warned – so the operator's strategy can't be to simply react to the display state; some form of internal memory for the task state is required.

How should the memory for the task state be represented in the architecture? Consider its properties: It must hold considerable information for many minutes at a time. Of the two conventional forms of memory in cognitive psychology, verbal working memory (VWM) or short-term memory (STM) is ruled out by the quantity of information, and also because the task requires substantial overt vocalization of interfering information. That is, the rehearsal process involved in conventional VWM or STM occupies the vocal processor and auditory memory. But in this task, the human operator has to speak and listen to many messages to other people, using track numbers, ranges, bearings, and other words that are also used in the to-be-remembered information. Conventional verbal WM/STM would simply be overwhelmed and would be effectively useless. The other conventional form of memory is long-term memory (LTM), but its normal characterization is that its write time is too slow (around 10 s/chunk) although retrieval from it can be fast.

GLEAN addresses this need with an explicit Task Memory component that stores properties of objects, optionally in a list structure, and has moderately slow read/write times, based on the GOMS general Mental Operator time of 1.2 s. The GLEAN models were consistent with task performance data, so the time parameters are reasonable. GLEAN's Task Memory was designed to be roughly consistent with Ericsson and Kintsch's (1995; Ericsson & Delaney, 1999) concept of skilled memory, termed Long-term Working Memory (LTWM). This has much faster store times than ordinary LTM, and much greater capacity than STM; the disadvantage is that it requires a task-specific "retrieval structure" that takes considerable practice in an information domain is develop. But at this time we lack a good theoretical proposal for how LTWM should be represented in terms of a comprehensive cognitive architecture. GLEAN's mechanism is simply a convenient ad-hoc solution.

The goal in this work was to first determine from a naïve model how much task state information needed to be stored over what time period. The result with the current EPIC model is that like the GLEAN model, quite a bit of such information is needed to avoid duplicating track actions. The next step would be to determine whether more sophisticated task strategies could reduce the amount of stored information. For example, under some conditions (e.g. low event rate) it might be possible to avoid storing the information for a long time by assuming that the track was queried when it should have been. At this time, the speculation is that this might help reduce the amount of stored information, but a significant amount will still be required.

### Modeling remote robot-control interfaces

A departure from planned work was work with Jill Drury and Jean Scholtz, well-known researchers in Human-Robot Interaction (HRI), on the prospects of using GOMS models to analyze the user interfaces for controlling search-and-rescue (SAR) robots. GOMS models in this domain are especially useful because of the great difficulty and cost of setting up user testing. his is another example of a complex task where modeling can be applied to achieve practical benefits.

## References


Ericsson, K.A., & Delaney, P.F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control.* New York: Cambridge Univerity Press. 257-297.

Ericsson, K.A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102 (2), 211-245.

Findlay, J. (1997). Saccade target selection during visual search. *Vision Research*, 37, 617–631.

Kieras, D.E, & Marshall, S.P. (2006). Visual Availability and Fixation Memory in Modeling Visual Search using the EPIC Architecture. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, 423-428.

Osga, G., Van Orden K., Campbell, N., Kellmeyer, D., and Lulue D. (2002). Design and Evaluation of Warfighter Task Support Methods in a Multi-Modal Watchstation. Space & Naval Warfare Center, San Diego, Tech Report 1874.

Peterson, M.S., Kramer, A.F., Ranxiao, F.W., Irwin, D.E., & McCarley, J.S. (2001). Visual search has memory. *Psychological Science*, 12, 287-292).

Rossetti, Y, & Pisella, L. (2002). Several 'vision for action' systems: a guide to dissociating and integrating dorsal and ventral functions. In W. Prinz & B. Hommel (Eds) *Common Mechanisms in Perception and Action: Attention and Performance XIX.* Pp. 62-119. Oxford: Oxford University Press.

Santoro, T.P, Kieras, D., & Pharmer, J. (2003). Verification and validation of latency and workload predictions for a team of humans by a team of computational models. (Tech. Rep. No. TR1227). Groton, CT, Naval Submarine Medical Research Laboratory. May 1, 2003.

St. John, M., Marshall, S. P., Knust, S. R., & Binning, K. R. (2006). Attention Management on a Geographical Display: Minimizing Distraction by Decluttering (Technical Report No. 06-01). San Diego, CA: San Diego State University, Cognitive Ergonomics Research Facility.

Williams, L.G. (1967). The effects of target specification on objects fixated during visual search. In A.F. Sanders (Ed.) *Attention and Performance*, North-Holland. 355-360.

# Products

### Graduate Student Statistics

Consistent with the project budget, there were no graduate students supported by this project.

### Journal Articles

Kieras, D. (in press). The persistent visual store as the locus of fixation memory in visual search tasks. *Cognitive Systems Research.*

### Book Chapters

Drury, J.L., Scholtz, J., Kieras, D. (2007). The Potential for Modeling Human-Robot Interaction with GOMS. In A. Lazinica (Ed.), *Human-robot interaction.* Advanced Robotic Systems International, Vienna, Austria, EU. ISBN 978-3-902613-13-4

Kieras, D.E. (2007). The control of cognition. In W. Gray (Ed.), *Integrated models of cognitive systems.* (pp. 327 - 355). Oxford University Press.

Kieras, D.E. (2008). Model-based evaluation. In J. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook* (2nd Ed). New York: Lawrence Erlbaum Associates/CRC Press. pp. 1191-1208.

Kieras, D.E. (in press). Model-based evaluation. In J. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook* (3rd Ed). New York: Taylor & Francis.

### Invited Presentations

Kieras, D.E. (2006). Thinking more broadly about cognitive architecture: Lessons from computer technology. Keynote talk presented at the AAAI Spring Symposium: Between a Rock and a Hard Place: Cognitive Science Principles Meet AI-Hard Problems. Stanford University, March 27-29, 2006.

Kieras, D.E. (2006). Complex Models + Complex Tasks = Complex Validation. Presented at the AFOSR Workshop on Model Comparison and Model Validation. Syracuse, N.Y. Sept. 7-9.

Kieras, D. Revisiting Vision: Unconventional Wisdom. Invited symposium presented at Rensselaer Polytechnic Institute, August 29, 2007.

Kieras. D. Revisiting Vision: Unconventional Wisdom. Invited symposium presented at Naval Research Laboratory, Center for Applied Research in Artificial Intelligence, Nov. 5, 2007.

### Refereed Conference Presentations

Kieras, D., & Knudsen, K. (2006). Comprehensive Computational GOMS Modeling with GLEAN. In *Proceedings of BRIMS 2006*, Baltimore, May 16-18, 2006. 11 pp.

Kieras, D.E, & Marshall, S.P. (2006). Visual Availability and Fixation Memory in Modeling Visual Search using the EPIC Architecture. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, 423-428.

Butler, K.A., Zhang, J., Esposito, C., Bahrami, A., Hebron, R, & Kieras, D. (2007). Work-centered design: A case study of a mixed-initiative scheduler. In *Proceedings of CHI 2007: Human Factors in Computing Systems.* New York: ACM, Inc. 747-756.

Drury, J. L., Scholtz, J., and Kieras, D. E. (2007). Adapting GOMS to Model Human-Robot Interaction. In *Proceedings of the Second Annual Conference on Human-Robot Interaction*, Arlington, VA, March 2007.

Douglass-Olberg, C/, Farkas, D., Steehouder, M., Karreman, J., Kieras, D., Roesler, A., Dalal, N., Baker, R., and Brunet, D. (2008): The new face of procedural content: a real world approach. In *Proceedings of ACM CHI 2008 Conference on Human Factors in Computing Systems*, April 5-10, 2008. pp. 3495-3500.

Kieras, D. (2009). Why EPIC was Wrong about Motor Feature Programming. In A. Howes, D. Peebles, R. Cooper (Eds.), Proceedings of *The 9th International Conference on Cognitive Modeling – ICCM2009*, Manchester, UK.

Kieras, D. (2009). The persistent visual store as the locus of fixation memory in visual search tasks. In A. Howes, D. Peebles, R. Cooper (Eds.), Proceedings of *The 9th International Conference on Cognitive Modeling – ICCM2009*, Manchester, UK.

Kieras, D. (2010). Modeling Visual Search of Displays of Many Objects: The Role of Differential Acuity and Fixation Memory. In *Proceedings of the 10th International Conference on Cognitive Modeling – ICCM2010*, Philadelphia, PA.

### *Technical Reports*

Drury, J. L., Scholtz, J., and Kieras, D. E. (December, 2006). Modeling Human-Robot Interaction with GOMS. MITRE Technical Paper. December 2006. Retrievable from http://www.mitre.org/work/tech_papers/tech_papers_06/06_1289/index.html

### *Awards/Honors*

*Jack A. Kraft Innovator Award*, presented by the Human Factors and Ergonomics Society, September 2010, "in recognition of significant efforts to diversify and extend the application of human factors principles to new areas of endeavor."

*Election to the SIGCHI Academy* by the ACM Special Interest Group for Computer-Human Interaction, April 2010, "for leadership in the profession of Computer-Human Interaction."

## Technology Transfer

*Background:* Starting with an initial contact in 1997 with NSWC Dahlgren, I began working with Tom Santoro of NSMRL to develop and apply Glean models to analyze the design of the Multi-Modal Watch Station for application to anti-air warfare tasks as part of the SC-21 Manning Affordability initiative project, which resulted in several conference publications and a journal article. This collaboration continued with Glenn Osga's group at SPAWAR, and also took on a new domain, the Land Attack Tomahawk control system prototype developed at SPAWAR.

In 2003, Soar Technology, Inc. licensed the GLEAN and EPIC intellectual property from the University of Michigan, and started a Phase I SBIR in collaboration with me (as consultant) to develop a commercialized version of GLEAN that also extended to designing for human error, led by Scott Wood, to follow up on his work on modeling error-handling in GOMS models. This led to a Phase II SBIR currently underway at Soar Technology; this work was described in a paper presented at the 2006 BRIMS conference, and is currently using as a modeling testbed the Tactical Tomahawk Weapon Control System (TTWCS) being developed by Lockheed-Martin. A group at MITRE working on modeling of air traffic control problems has also starting using GLEAN - see the brief description in section 2d. With Soar Technology, I joined a group led by Glenn Osga of SPAWAR in a Capable Manpower proposal on advanced user interface development techniques, where GLEAN could contribute to the analysis

I continued as a consultant working on an SBIR Phase II project to produce a commercial version of GLEAN with Soar Technology, Inc., using my software licensed from University of Michigan, an effort led by Scott Wood, now based in a new startup company, Highly Human Software, that now has the licenses to commercialize GLEAN and EPIC. The modeling testbed in this project is the Tactical Tomahawk Weapon Control System (TTWCS) being developed by Lockheed-Martin.

Steven Estes and his group at MITRE working on modeling of air traffic control systems has been using GLEAN under licensing, both doing model development

Both GLEAN and EPIC were involved in a project conducted with Keith Butler at The Boeing Company during 2004-2005 that applied modeling techniques to the design of an advanced maintenance/operations scheduling application for the U.S. Air Force. GLEAN was used in a novel technique, high-level GOMS models, to assist in the function-allocation problem, and an EPIC model was used to verify the feasibility of detailed modeling of a complex display-based user procedure. Together with other techniques, the result was a high-productivity system whose extremely effective user interface was designed in very little time, basically with only one user-testing iteration. This work was presented in a paper at the ACM CHI 2007 conference. Additional work was done under a consulting arrangement with Microsoft, where Keith Butler relocated.

Steven Estes and his group at MITRE working on modeling of air traffic control systems has been using GLEAN under licensing, both doing model development with the commercial prototype developed at Soar Technology, and the full-scale simulations with the "kernel" – my basic GLEAN simulation engine which has been plugged into their large scale modeling environment. I continue to be impressed with the large size of the models they have constructed, which validates the basic architecture and software approach, shared with EPIC.

Another collaboration concerns much older ONR-sponsored research on procedural text and documentation. I served as a consultant to a group at Microsoft who are re-engineering the technical support web pages for key user problems. The input on troubleshooting, mental models, and effective procedural text construction that I developed under previous ONR projects was combined with other concepts to produce much more effective support materials. A preliminary report on this work was presented at the 2008 ACM CHI conference:

Other technology transfer activities involved other researchers who are using EPIC and GLEAN. These include Derek Brock's group at NRL, Anthony Hornof at University of Oregon, and Travis Seymour at University of California-Santa Cruz, who is also using EPIC to teach cognitive modeling. In addition, Scott Wood, currently with the VA's patient safety group, is using GLEAN to model safety-critical aspects of the VA's main electronic patient records system.

## Final Report Distribution List

Paul Bello
Office of Naval Research
875 N. Randolph St.
Arlington, VA 22203-1995

Office of Naval Research
Regional Office - Chicago-N62880
230 South Dearborn, Room 380
Chicago IL 60604-1595

Defense Technical Information Center
8725 John J. Kingman Road STE 0944
Fort Belvoir, VA 22060-6218

Naval Research Laboratory
Attn: Code 5596
4555 Overlook Avenue SW
Washington, DC 20375-5320